Mol Gen Genet (1996) 251:52–59

ORIGINAL PAPER

Susan W. Gorman · Dawn Banasiak
Cecelia Fairley · Sheila McCormick

A 610 kb YAC clone harbors 7 cM of tomato (*Lycopersicon esculentum*) DNA that includes the *male sterile 14* gene and a hotspot for recombination

Received: 6 July 1995 / Accepted: 8 November 1995

Abstract Pollen development requires both sporophytic and gametophytic gene expression. We are using a map-based cloning technique to isolate sporophytic genes which, when mutant, cause pollen abortion and a male sterile (*ms*) phenotype in tomato (*Lycopersicon esculentum*). We have genetically characterized one *ms* locus (*ms14*) using RFLP analysis and identified flanking markers. High-resolution genomic physical mapping indicates that the *ms14* locus is located in a ~300 kb region. We have identified a YAC clone with an insert size of ~610 kb that contains the *ms14*-linked markers, reflects the organization of the physical map and therefore most probably contains the *ms14* gene. In addition, we present evidence that the relationship between physical and genetic distance in this chromosomal region changes abruptly from ~105-140 kb/cM to less than 24 kb/cM, and suggest that the TG393-TG104 region is a hotspot for recombination.

Key words Positional (map-based) cloning · Pulsed field electrophoresis · Physical mapping · RFLP mapping · Pollen development

Introduction

Pollen development in higher plants is a complex process that requires the coordination of gene expression in both the sporophytic and gametophytic tissues of the anther (reviewed in McCormick 1993). Mutations in male sterile (*ms*) loci result in aborted pollen and generally require the homozygous recessive condition in order to exhibit the phenotype. This suggests that these genes act in the sporophytic tissues of the anther but affect the developing microspores (gametophytes). Male sterile genes might be expressed before or during meiosis in the sporophytic pollen mother cell or in the tapetum or other sporophytic cells of the anther. We hope to develop an understanding of the sporophytic contribution to pollen development by studying the products of male sterile loci.

Current biochemical and molecular information do not provide a clear picture of the mechanism(s) of nuclearly encoded male sterility. Although three genes have been isolated that produce a male sterile phenotype when mutant, their gene products do not provide obvious clues to the roles that these genes play during pollen development. For example, Moffatt and Somerville (1988) isolated an adenine phosphoribosyl transferase (*aprt*$^-$) *Arabidopsis* mutant, but how a lesion in a nucleic acid salvage pathway disrupts pollen development is unclear. The *ms2* gene from *Arabidopsis* (Aarts et al. 1993) contains a small region homologous to a wheat mitochrondrial ORF, but no function has been ascribed to this mitochondrial region or to the gene product of the *ms2* gene. In maize, a male sterile gene designated *Ms45* (M. Albertsen, personal communication) was cloned (Albertsen et al. 1993) that showed 33% homology over 256 amino acids with strictosidine synthase. Although the action of strictosidine synthase in the indole alkaloid pathway is understood, it is not clear whether or how indole sythesis might be involved in pollen development. These examples serve to illustrate the complexities of male gametophyte development and, therefore, it will probably be necessary to isolate and characterize a large number of male sterile gene products before we can develop a comprehensive understanding of the sporophytic contribution(s) to pollen development. In tomato (*Lycopersicon esculentum*), more than 40 spontaneous *ms* loci have been identified (summarized in Kaul 1988). We have targeted for isolation one of these male sterile genes, *ms14*; in

Communicated by R. Hagemann

S. W. Gorman · D. Banasiak · C. Fairley · S. McCormick (✉)
Plant Gene Expression Center, USDA/ARS and University of California-Berkeley, 800 Buchanan St., Albany, CA 94710, USA

ms14 mutants pollen development is blocked after the first microspore mitosis (Rick 1948).

Although the time of phenotypic arrest is known for most of the tomato *ms* mutants, the function of the corresponding genes is unknown and no biochemical assays exist for their gene products. Map-based cloning in tomato is a workable strategy for isolation of these genes since the technique requires only that the target gene have a clear phenotype and that its position on a genetic map be known. Map-based cloning has been used successfully to isolate target genes from *Arabidopsis* (Giraudat et al. 1992; Arondel et al. 1992; Bent et al. 1994; Putterill et al. 1995) and tomato (Martin et al. 1993a). This technique involves identifying markers closely linked to the target locus, determining the physical/genetic relationship (kb/cM) in the target region, isolating clones that contain the locus from a large insert library and confirming the locus identity by complementation of the recessive phenotype with the dominant allele. *ms14* is an attractive candidate for map-based cloning because it is located on the classical map of chromosome 11 close to the I_2 gene; I_2 is linked to TG105 on the RFLP map of chromosome 11 (Sarfatti et al. 1989).

Of the steps involved in positional cloning, constructing and integrating detailed genetic and physical maps for the region of interest are arguably the most important. Although a chromosome walk can be initiated after identification of closely linked genetic markers, establishing the relationship between physical and genetic distance for the locus of interest can help predict the feasibility of the walk. The values for the average physical/genetic distance in tomato range from ~750 kb/cM (Tanksley et al. 1992) to ~900 kb/cM (Patterson and Wing 1993). However, specific chromosomal regions are known to diverge widely from these averages; for example, values from ~4–16 Mb/cM have been reported for the region at the *Tm2a* locus (chromosome 9; Ganal et al. 1989b),while 43 kb/cM was determined for a region that is centromere distal to the I_2 locus (chromosome 11; Segal et al. 1992). Until recently it was not known whether this variation occurred gradually over some distance or changed rapidly between adjacent chromosomal regions. Civardi et al. (1994) presented evidence for rapid change in a small chromosomal region using the maize *a1* and *sh2* loci. They found that the region between these loci had a physical/genetic relationship of 1560 kb/cM, while the value for a 1 kb interval within the *a1* gene was only 217 kb/cM. These results suggest that the genome may be punctuated with areas of increased recombination that lead to abrupt changes in the physical/genetic relationship.

In this paper we present genetic and physical mapping data indicating that *ms14* lies within a 300 kb region defined by the RFLP markers CT120A and TG393. Using these markers, we provide molecular evidence indicating that a 610 kb YAC clone of genomic DNA harbors the *ms14* locus. We also show that the physical/genetic relationship in the chromosomal region containing the *ms14*-linked markers changes abruptly and suggest that the TG393-TG104 region is a hotspot for recombination.

Materials and methods

Plant materials

The *ms14* mutation (Rick 1948) was identified as a spontaneous mutant in the *L. esculentum* cultivar Earliana. To generate mapping populations for the male sterile mutant, the *ms/ms* plants were identified in a *ms/ms* × *Ms/ms* segregating population and used as female parents in a cross with *L. pennellii* (LA716, a self-compatible accession). Crosses between *L. esculentum* and *L. pennellii* succeed only when *L. pennellii* serves as the male parent. The F_1 progeny are fertile, but there are fertility problems with the F_2 of such interspecies crosses (C. M. Rick, personal communication), making it difficult to score accurately for the presence of *ms* genes. To circumvent this problem, we generated a backcross mapping population by crossing F_1 plants (*Ms/ms*) as male parents to *ms/ms L. esculentum* females. Such interspecific backcross progeny have no problems with fertility and are only male sterile if they are genetically *ms/ms* at a male sterile locus. Backcross progeny were scored for fertility after squashing the anthers from a flower in a microtiter well that contained 100 µl of pollen germination medium (Jahnen et al. 1989) and examining the anther contents under a dissecting microscope. Wild-type pollen appeared round with a diameter of ~25–30 µm but *ms14* pollen had a diameter of only ~10 µm. The plants segregated 1:1 for male sterility and each male sterile plant was rescored multiple times on subsequent dates. Using a pooled-DNA sample technique (Churchill et al. 1993), recombinant plants were identified in the mapping population of 220 male sterile plants.

Molecular probes and recombinant DNA methodology

Enzyme survey blots (ESB; Bernatzky and Tanksley 1986) were prepared from DNA extracted (Bernatzky and Tanksley 1986) from *L. esculentum* cv. VF36 and *L. pennellii* (LA 716) for analysis of polymorphism. Standard gel electrophoresis was conducted using 1% agarose, 250 ng/ml ethidium bromide, 1 × TAE gels. DNA was transferred from the gels to Nytran membranes (Schleicher and Schuell) as described by Southern (1975).

Chromosome 11 RFLP clones (Tanksley et al. 1992) were obtained from S. D. Tanksley (Cornell University). RFLP inserts were sequenced with Sequenase (USB) using the forward or reverse primers from the vector. The partial sequences have been deposited in Genbank. Appropriate PCR primers were designed and amplification reactions of the inserts carried out using 100 mM dNTP, 1 unit Taq polymerase, 10 mM DTT, 50 mM KCl, 10 mM TRIS-HCl (pH 8.3), 1.5 mM $MgCl_2$, 0.01% Tween-20, 0.01% NP-40, 20 pmol of each primer and 1 µl (~10 ng of clone or 50 ng of genomic miniprep) DNA. Reaction conditions were 3 min at 95°, then 35 cycles of 1 min at 94°, 1 min at 58° and 1 min at 72°. PCR products were separated on 1.5% agarose gels, bands isolated and DNA recovered using Geneclean (Bio 101) or Qiaex (Qiagen). RFLP marker inserts were labeled with [α-^{32}P]dATP using the method of Feinberg and Vogelstein (1984).

Pulsed field gel electrophoresis

The second, third and fourth sets of leaves from 10-15 cm tall *L. esculentum* (cv. VF36) plants were used to prepare protoplasts. DNA

was prepared for CHEF (clamped-contour *h*omogeneous *e*lectrophoretic *f*ield) gels by encapsulating and lysing protoplasts in agarose (Ganal and Tanksley 1989a), cleaving with rare-cutter restriction enzymes and separating the products on gels using a Bio-Rad CHEF DR-III Megabase DNA pulsed field electrophoresis system (Bio-Rad Laboratories). Initial CHEF gels were run under conditions which separated fragments ranging from 200 kb to 2.2 Mb. Based on these results and the predictions for ensuing digests, subsequent separations were optimized for narrower separation ranges (e.g. 50–500 kb, 500–1000 kb, 1000–4000 kb, etc.) in order to estimate band sizes more accurately. Gels were incubated in 1 mg/ml ethidium bromide for 30 min and the DNA nicked by exposure to 60 mJ of UV in a Stratalinker (Stratagene). DNA was transferred to Nytran membrane (Schleicher and Schuell) using the method of Southern (1975) and fixed to the membrane with 120 mJ of energy in a Stratalinker. To increase accuracy, some fragments (sized 50 kb or less) were also sized on 0.3% agarose gels using conventional electrophoresis. For sequential probings, blots were stripped of radioactivity by immersion in boiling 0.1 × SSPE, 0.5% SDS for 20 min.

Pulsed field gel electrophoresis and hybridizations of YACs

YAC inserts were sized using a Bio-Rad CHEF DR-III Megabase DNA pulsed field electrophoresis system (Bio-Rad Laboratories) optimized for 200 kb to 2.2 Mb separations. Briefly, cells of a single YAC clone were encapsulated in agarose plugs and incubated at 37° for 2.5 h in 1 M sorbitol, 20 mM EDTA, 14 mM DTT, 10 mM TRIS-HCl, pH 7.5 and 1 mg/ml zymolase. Plugs were transferred to 1% lithium dodecyl sulfate, 100 mM EDTA, 10 mM TRIS-HCl, pH 8.0 and lysed overnight at 37°. Plugs were washed three times for 30 min at 50° and for 30 min at room temperature prior to electrophoresis. Sizes were estimated by comparison with a standard curve generated from the migration distances of *Saccharomyces cerevisiae* chromosomes and/or λ DNA concatamer controls.

Mapping gels were constructed by cleaving the agarose-encapsulated DNA with 40 units of rare-cutter restriction enzyme in a total volume of 250 µl for 1 h and separating the products on CHEF gels containing λ DNA concatamer controls. After electrophoresis, gels were transferred and prepared for hybridization as described for genomic CHEFs. Membranes containing products from partial *Aat*I, *Bss*HII, *Eag*I, *Fsp*I, *Mlu*I, *Nae*I, *Not*I, *Nru*I, *Pvu*I, *Sal*I, *Sfi*I, *Sma*I and *Xho*I digests were hybridized with probes specific to the left or right arms of the YAC vector and CT65, CT120A, TG393 and TG104. For sequential probings, blots were stripped of radioactivity by immersion in boiling 0.1 × SSPE, 0.5% SDS for 20 min.

Results

Genetic mapping of *ms14*

The *ms14* locus, located at position 88 on the chromosome 11 classical map, was mapped relative to the *anthocyaninless* locus (*a*, position 68) and the *hairless* locus (*hl*, position 48) (Clayberg 1970; Tanksley et al. 1992; see Fig. 1). Since the proximal I_2 gene (located at map position 85; Sarfatti et al. 1989) was reported to be tightly linked to RFLP TG105, we restricted linkage analysis to those RFLP markers located in the region between TG105 and the telomere of chromosome 11. Single plants from a backcrossed population were scored for RFLPs TG26, CT65, CT120A, TG393, TG104 and for male sterility. This initial analysis placed the *ms14* locus between RFLPs CT120A and TG393/TG104. We then used pooled-sample mapping (see Churchill et al. 1993; Giovannoni et al. 1995) on a backcrossed population of 220 male sterile plants to assess genetic distances between the CT120A, TG393, TG104 and *ms14* loci. Table 1 contains the recombination fractions obtained for these markers. Pools displaying recombination between *ms14* and a RFLP marker were subjected to single plant analysis and the recombinant individual(s) identified. We found *ms14* to be about equidistant between CT120A and TG393. CT120A was 2.3 cM from TG393 and, in this population, TG393 was separated from TG104 by ~2.3 cM.

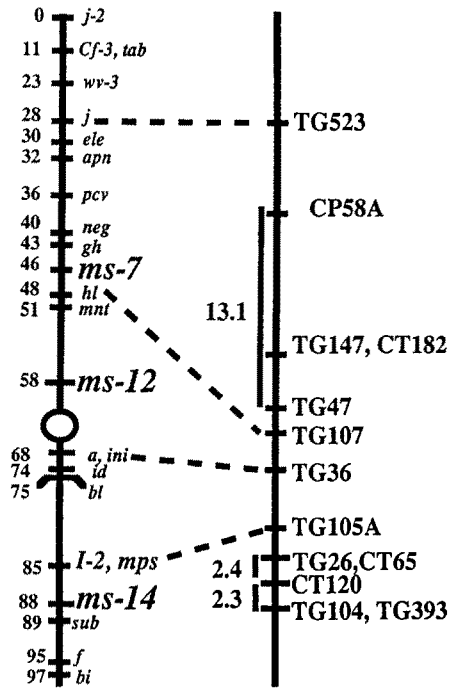

Fig. 1 Interdigitation of classical and RFLP maps of chromosome 11. The classical map positions of morphological markers are located on the left. The RFLP map is on the right and recombination frequencies reported by Tanksley et al. (1992) appear next to a bar connecting selected RFLPs. The linkage relationships between the classical and RFLP maps as reported by Tanksley et al. (1992) are indicated by *dotted lines*

Creating physical maps of the *ms14* region

In order to determine the relationship between physical and genetic distance for the chromosomal region containing the *ms14*-linked RFLPs, we constructed physical maps for these loci. To do this, megabase sized DNA was isolated from *L. esculentum* (cv.VF36) protoplasts, cleaved with methylation-sensitive enzymes having a 6–8 bp recognition sequence, separated via pulsed-field electrophoresis (CHEF) and analysed by Southern

Table 1 Recombination fractions for markers linked to *ms14*

Markers	Recombination fraction
CT120A-TG393	0.0227 ± 0.0100
TG393-TG104	0.0227 ± 0.0100
CT120A-TG104	0.0363 ± 0.0126
CT120A-*ms14*	0.0136 ± 0.0078
TG393-*ms14*	0.0181 ± 0.0090
TG104-*ms14*	0.0363 ± 0.0126

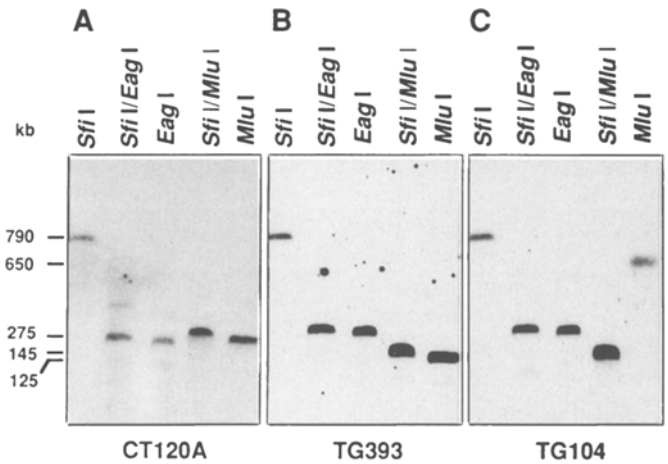

Fig. 2A–C Hybridization of a tomato genomic CHEF gel blot with three *ms14*-linked RFLP markers. Agarose-encapsulated high molecular weight DNA was digested and separated on a CHEF gel optimized for separation in the 500–1000 kb range, Southern blotted and sequentially probed with (**A**) CT120A, (**B**) TG393 and (**C**) TG104. Sizes of the hybridizing bands appear to the left

blotting. To identify bands hybridizing to more than one RFLP marker (Fig. 2), RFLPs were sequentially hybridized to at least two blots prepared from independent gel runs of the same DNA preparation.

Figure 3A depicts the physical maps generated for the CT120A, TG393 and TG104 loci, using data obtained from single and double digests. We began by cleaving megabase sized DNA with the enzymes *Sal*I, *Eag*I, *Fsp*I, *Mlu*I and *Nru*I. A single *Eag*I band of 275 kb was detected when this blot was hybridized sequentially with the RFLP markers TG26, CT120A, TG104 and TG393 (Fig. 2 and data not shown), suggesting that the distance between TG26 and TG104 was no more than 275 kb. This predicted a maximum value for the physical/genetic relationship of ~40 kb/cM for this region, a value similar to the 43 kb/cM found by Segal et al. (1992) for the chromosome 11 region distal to the I_2 locus. However, double digests using the enzymes *Sal*I, *Fsp*I, *Mlu*I, *Nae*I and *Nru*I (data not shown) indicated the presence of three co-migrating *Eag*I fragments, one carrying RFLP TG26, one carrying CT120A and one carrying both TG393 and TG104.

To establish physical linkage between the maps for CT120A, TG393 and TG104, we identified enzymes generating larger fragments that hybridized with all three *ms14*-linked RFLPs. Hybridization of the *ms14*-linked markers to high molecular weight DNA digested with *Bss*HII generated a fragment of 950 kb, while *Sfi*I generated a fragment of 790 kb, suggesting that CT120A, TG393 and TG104 were located within a 790 kb region (Fig. 2 and data not shown). Digestion with *Bss*HII or *Sfi*I and additional rare-cutter enzymes (*Mlu*I, *Eag*I, *Sal*I, *Pvu*I, *Nae*I, and *Xho*I), produced fragments whose combined molecular weights were less than those of the *Bss*HII or *Sfi*I fragments (Fig. 2 and data not shown). This result is consistent with the hypothesis that CT120A, TG393 and TG104 are located within a 790 kb region. The physical map derived for this region is depicted in Fig. 3B. Using this map, we calculated a distance of ~26 kb between TG393 and TG104 and ~280 kb between CT120A

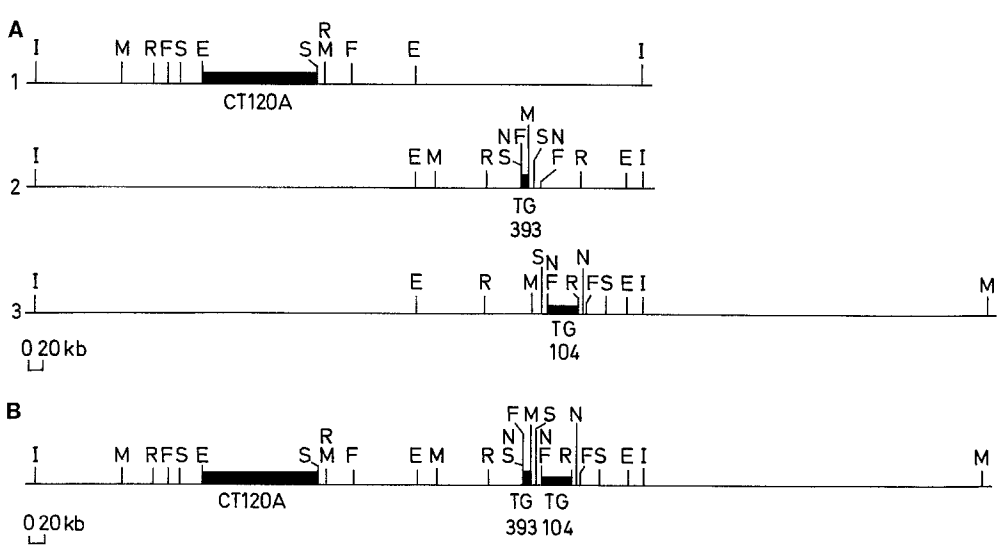

Fig. 3A, B Physical maps of the tomato chromosome 11 region containing the *ms14*-linked RFLP markers. **A** Maps constructed for individual RFLPs CT120A (1), TG393 (2) and TG104 (3). The RFLP marker is located within the region represented by a *black bar*. **B** Alignment of the three RFLP maps with respect to one another. The enzymes *Eag*I (E), *Fsp*I (F), *Mlu*I (M), *Nae*I (N), *Nru*I (R), *Sal*I (S), and *Sfi*I (I) were used

and TG393. Comparison to the genetic map yields a value of ~12 kb/cM for the region between TG104 and TG393 and of ~123 kb/cM for the region between CT120A and TG393.

Using the maps for chromosome walking predictions

The relatively short physical distance between the *ms14* flanking markers makes a chromosome walk to the *ms14* locus feasible. We obtained a single YAC clone known to contain the TG104 marker from the Keygene (Wageningen, The Netherlands) tomato YAC library. This clone, YAC 2/279, contained an insert of ~610 kb. We constructed a map for this clone after digesting YAC DNA with a total of 13 rare cutter enzymes (six are shown in Fig. 4) and hybridizing with the *ms14*-linked markers. TG104 is located within 75 kb of the right end of the YAC, TG393 was located at most 55 kb from TG104, and CT120A maps within an area 238-318 kb distant from the region containing TG393 (Fig. 4). These results are consistent with the genomic physical map generated for the CT120A, TG393 and TG104 chromosomal region. In addition, we found that the marker CT65 maps to a region ~168–247 kb from CT120A and within ~24 kb of the left end of the YAC (Fig. 4). As further confirmation for these marker placements, we digested both YAC and genomic DNA with *Dra*I, *Eco*RI, *Hae*III, *Kpn*I and *Xba*I, and hybridized the Southern blots with the four RFLP markers located on the YAC. We found no difference in size between the YAC and genomic fragments, supporting the conclusion that the YAC insert correctly reflects the genomic organization of these fragments (data not shown).

The YAC map also confirms the estimates of the physical/genetic relationship between markers in this region. Depending on the placement of the markers within the hybridizing fragment, we obtain a value of 105–140 kb/cM for the region containing *ms14* and bounded by CT120A and TG393, in keeping with the estimate of ~123 kb/cM predicted from the genomic physical map analysis. The genomic physical map value of ~12 kb/cM for the TG393/TG104 region also fell within the range determined from the YAC map, where the maximum value was ~24 kb/cM. Hence the physical/genetic relationship in this region is at least 4.4 times less than that in the adjacent CT120A/TG393 region, indicating a possible recombinational hotspot. Although we did not include the marker CT65 in our pooled-sample mapping or our genomic physical mapping analysis, we can estimate the physical/genetic relationship for the region between CT65 and CT120A at about 70-103 kb/cM from the YAC map, using the Tanksley et al. (1992) value of 2.4 cM for the genetic distance between these markers.

Discussion

We are focusing our efforts on isolating male sterile genes from spontaneous mutants using a map-based cloning approach. Here we report the genetic and physical characterization of the tomato chromosomal region that harbors one such gene, *ms14*, and the isolation of this region on a YAC. In addition, we show that the relationship between physical and genetic distance changes abruptly in this region and propose that the region between two *ms14*-linked RFLP markers is a recombinational hotspot.

Genetic maps

Our goal is to clone the gene responsible for the *ms14* phenotype in tomato by map-based cloning. Although RFLP and classical genetic maps were available for all the tomato chromosomes when this study was initiated

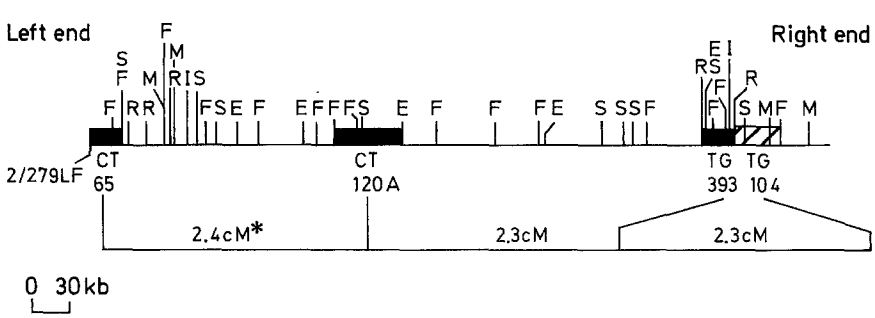

Fig. 4 Physical map of YAC 2/279. Agarose-encapsulated chromosomes were partially digested with rare cutter restriction enzymes, separated on CHEF gels and analyzed by Southern blotting. A map was constructed after sequential hybridization with probes specific for the left and right ends of the YAC vector. RFLP markers CT65, CT120A, TG393 and TG104 were subsequently hybridized to the blots and placed on the map. The RFLP markers are located within the regions represented by the *black bars*. The genetic distance between markers (in cM, drawn to scale) appears below the physical map. The *asterisk* indicates the genetic distance obtained by Tanksley et al. (1992); all other values were determined in this study. The enzymes *Eag*I (E), *Fsp*I (F), *Mlu*I (M), *Nru*I (R), *Sal*I (S) and *Sfi*I (I) were used

(Zamir and Tanksley 1988), only one link between the morphological and RFLP maps of chromosome 11 had been reported (Sarfatti et al. 1989). Since RFLP maps and classical maps are based on different populations and the map positions of individual markers can differ considerably in statistical accuracy, analysis of RFLP and phenotype segregation must be done in the same population to find those molecular markers that are most closely linked to the phenotype (see Weide et al. 1993; Balint-Kurti et al. 1995; Giovannoni et al. 1995). We find that *ms14* is located in the middle of the 2.3 cM region between CT120A and TG393 using pooled-sample analysis (Churchill et al. 1993) on a backcrossed population of 220 male sterile plants.

Physical maps

We constructed physical maps of the RFLP loci that flank the *ms* genes in order to calculate the average physical/genetic relationship for this chromosomal region and to estimate the feasibility of a chromosome walk without additional markers. Our mapping results suggest that single digests can be misleading due to the presence of co-migrating fragments. We found that what initially appeared to be one 275 kb *Eag*I fragment carrying four *ms14*-linked markers was actually three different 275 kb *Eag*I fragments: one carries TG26, one has CT120A and one contains TG393 and TG104. Wing et al. (1994) also report co-migration of fragments (175 kb *Mlu*I) bearing markers flanking the *jointless* locus which were detected after double digests were conducted. The general assumption has been that physical linkage is indicated by hybridization of linked loci to large fragments produced from single digests. In view of our findings and those of Wing et al. (1994), it would seem that double digest analysis can avoid the risk of miscalculating the average physical/genetic relationship for a chromosomal region and better estimate the feasibility of a chromosome walk. Using this approach, we determined that the *ms14* gene is located in the middle of the ~300 kb region bounded by CT120A and TG393, making a chromosomal walk to the gene from either flanking marker feasible.

Since the enzymes we used for the genomic physical mapping are all methylation sensitive, the patterns of restriction sites represent the methylation state of the genome for a particular developmental stage. We found a clustering of the methylation-sensitive restriction sites *Fsp*I, *Mlu*I, *Nae*I, *Nru*I, and *Sal*I associated with the TG393 and TG104 loci, suggestive of CpG-like islands (see Fig. 3A). CpG islands are defined as sequences of 0.2–3.0 kb with high GC levels and unmethylated CpG doublets (Bird 1987; Larsen et al. 1992). In mammals, CpG islands overlap the transcription start site of all housekeeping genes and include the transcription start site or other parts of the transcription unit in approximately 60% of genes that are expressed in a tissue-specific manner (Larsen et al. 1992). CpG islands are present in plants and although plants also possess methylated CpXpG sequences (unlike vertebrates where cytosine methylation is confined to CpG), unmethylated CpXpG sequences are not as frequently associated with CpG islands (Antequera and Bird 1988; Messeguer et al. 1991). Recently, Larsen et al. (1992) identified additional rare-cutter enzyme sites that contain all four bases (e.g., *Mlu*I which recognizes the sequence ACGCGT) in addition to sequences composed exclusively of C and G in known CpG islands and classified them with respect to the frequency of their appearance. Several of these sites are present around TG393 and TG104 and additional methylation-sensitive enzyme sites also cluster near these loci, suggesting that CpG islands are closely associated with them or, at the least, that this chromosomal region is undermethylated in leaves of the young seedlings used for this analysis.

The relationship between physical and genetic distance

The tomato genome is 950 megabases long (Arumuganathan and Earle 1991) and recent calculations of the average physical distance per cM in tomato give values of 750 kb/cM (Tanksley et al. 1992) and 900 kb/cM (Paterson and Wing 1993). The actual relationship between genetic and physical distance has been examined in several regions of the tomato genome. On chromosome 5, in the region around the *Pto* locus, Martin et al. (1993b) estimated ~1088 kb/cM between RFLP TG 538 and another marker about 0.4 cM away. Van Daelen et al. (1993) found three genetically inseparable markers that each showed completely different restriction maps with no fragments in common (some greater than 2000 kb in size) after digestion with nine different enzymes for the region around the *Mi* locus, located near the centromere on chromosome 6. In contrast to these results, Segal et al. (1992) report a value of only 43 kb/cM for the 4.1 cM region on chromosome 11 that is distal to the I_2 locus, while estimates of ~200 kb/cM were made for the regions containing the *rin*, *nor* (chromosomes 5 and 10, respectively; Giovannoni et al. 1995) and *jointless* (chromosome 11; Wing et al. 1994) loci.

We have suggested a maximum value of ~24 kb/cM between TG104 and TG393. This value is comparable to the 43 kb/cM suggested for the region between TG105 and TG26 (Segal et al. 1992), located about 8.6 cM from TG393. However, our analysis also predicts a kb/cM value at least 4.4 times greater for the chromosomal region between CT120A and TG393 which lies within the section bounded by the TG26 and TG393 loci (see Fig. 1). This suggests that the TG393-TG104 sector and the TG105-TG26 region may be "hot spots" with elevated recombination values. "Hot

spots" are thought to occur in euchromatic regions of the chromosome where actively transcribed genes are located, whereas in heterochromatic regions, such as those surrounding centromeres, recombination is believed to be suppressed. In tomato, studies seem to support this idea. Values in excess of 4 Mb/cM have been reported for genes located in centromeric regions (Ganal et al. 1989b; Van Daelen et al. 1993) while values for loci distant from these regions are lower by an order of magnitude or more (e.g. Giovannoni et al 1995; Wing et al. 1994; Segal et al. 1992). If "hot spots" are located in euchromatic regions, is their increased recombination rate due to a concentration of genes in the region or to other factors responsible for maintaining the euchromatic state? Data is not yet available that indicates whether recombination occurs only within the genes located in "hot spots" or whether intergenic regions can also serve as breakpoints for recombination. In plants, only one example of a region with a large kb/cM value that is juxtaposed to a region with a notably smaller value has been reported (Civardi et al. 1994). We have documented a second occurrence of significant change in the physical/genetic relationship within a small chromosomal region in a plant genome. This abrupt change has enabled us to isolate a single YAC (2/279) with an insert of ~610 kb that corresponds to ~7 cM in the genome.

ms14 gene identification

Our genetic and physical mapping indicates that the *ms14* gene is located on YAC 2/279 in the ~300 kb region between CT120A and TG393. We are currently attempting to localize the position of *ms14* precisely within this region by conducting a walk from CT120A to TG393 in a cosmid library prepared from YAC 2/279. We are concurrently screening a cDNA library constructed from anthers representing premeiotic to free binucleate microspores using YAC 2/279 as a probe. cDNA screening with YAC clones has been used successfully to identify target genes in *Arabidopsis* (Arondel et al. 1992; Chang et al. 1993) and tomato (Martin et al. 1993a).

Acknowledgements We thank Pieter de Vos and colleagues at Keygene for graciously screening their YAC library with TG104 for us and David Hantz for greenhouse care of the plants. We also thank Anne Bengtson, Kristen Daehler, John Forte, Paul Guth, Elise Grenier, Dennis Hsueh and David Wilson for assistance with RFLP mapping, Rima Kulikauskas for sequencing inserts of RFLP clones and David Neale and Mark Gorman for helpful discussions. This work was supported by United States Department of Agriculture grant 91-37300-6495. Anne Bengtson, Kristen Daehler and John Forte were partially supported by a United States Department of Agriculture high school teacher summer internship program, and David Wilson was partially supported by a USDA high school student summer research program.

References

Aarts MGM, Dirkse WG, Steikema WJ, Pereira A (1993) Transposon tagging of a male steility gene in *Arabidopsis*. Nature 363:715–717

Albertsen MC, Fox TW, Trimnell MR (1993) Tagging, cloning and characterizing a male fertility gene in maize. Amer J Bot Abst 80:16

Antequera F, Bird AP (1988) Unmethylated CpG islands associated with genes in higher plant DNA. EMBO J 7:2295–2299

Arondel V, Lemieux B, Hwang I, Gibson S, Goodman HN, Somerville CR (1992) Map-based cloning of gene controlling omega-3 fatty acid desaturation in *Arabidopsis*. Science 258:1353–1355

Arumuganathan K, Earle ED (1991) Nuclear DNA content of some important plant species. Plant Mol Biol Rep 9:208–218

Balint-Kurti J, Jones DA , Jones JDG (1995) Integration of the classical and RFLP linkage maps of the short arm of tomato chromosome 1. Theor Appl Genet 90:17–26

Bent AF, Kunkel BN, Dahlbeck D, Brown KL, Schmidt R, Giraudat J, Leung J, Staskawicz BJ (1994) *RPS2* of *Arabidopsis thaliana*: a leucine-rich repeat class of plant disease resistance genes. Science 265:1856–1859

Bernatzky R, Tanksley SD (1986) Toward a saturated linkage map in tomato based on isozymes and random cDNA sequences. Genetics 112:887–898

Bird AP (1987) CpG islands as gene markers in the vertebrate nucleus. Trends Genet 3:342–347

Chang C, Kwok SF, Bleecker AB, Meyerowitz EM (1993) *Arabidopsis* ethylene-response gene *ETR1*: similarity of product to two-component regulators. Science 262:539–545

Churchill GA, Giovannoni JJ, Tanksley SD (1993) Pooled-sampling makes high-resolution mapping practical with DNA markers. Proc Natl Acad Sci USA 90:16–20

Civardi L, Xia Y, Edwards KJ, Schnable PS, Nikolau BJ (1994) The relationship between genetic and physical distances in the cloned *a1-sh2* interval of the *Zea mays* L. genome. Proc Natl Acad Sci USA 91:8268–8272

Clayberg CD (1970) Mapping male-sterile genes on chromosome 11. Report Tomato Genet Coop 20:13

Feinberg AP, Vogelstein B (1984) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132:6–13

Ganal MW, Tanksley SD (1989a) Analysis of tomato DNA by pulsed field gel electrophoresis. Plant Mol Biol Rep 7:17–27

Ganal MW, Young ND, Tanksley SD (1989b) Pulsed field gel electrophoresis and physical mapping of large DNA fragments in the *Tm-2a* region of chromosome 9 in tomato. Mol Gen Genet 215:395–400

Giovannoni JJ, Noensie EN, Ruezinsky DM, Xianghuai L, Tracy SL, Ganal MW, Martin GB, Pillen K, Alpert K, Tanksley SD (1995) Molecular genetic analysis of the *ripening-inhibitor* and *non-ripening* loci of tomato: a first step in genetic map-based cloning of fruit ripening genes. Mol Gen Genet 248:195–206

Giraudat J, Hauge BM, Valon C, Smalle J, Parcy F, Goodman HM (1992) Isolation of the *Arabidopsis ABI3* gene by positional cloning. Plant Cell 4:1251–1261

Jahnen W, Lush WM, Clarke WM (1989) Inhibition of in vitro pollen tube growth by isolated S-glycoproteins of *Nicotiana alata*. Plant Cell 1:501–510

Kaul MLH (1988) Male sterility in higher plants. Springer-Verlag, New York

Larsen F, Gundersen G, Prydz H (1992) Choice of enzymes for mapping based on CpG islands in the human genome. Gene Anal Techniques 9:80–85

Martin GB, Brommonschenkel SH, Chunwonse J, Frary A, Ganal MW, Spivey R, Wu T, Earle ED, Tanksley SD (1993a) Map-based cloning of a protein kinase gene conferring disease resistance in tomato. Science 262:1432–1436

Martin GB, De Vincente MC, Tanksley SD (1993b) High-resolution linkage analysis and physical characterization of the *Pto* bacterial resistance locus in tomato. Molec Plant-Microbe Interact 6:26–34

McCormick S (1993) Male gametophyte development. Plant Cell 5:1265–1275

Messeguer R, Ganal MW, Steffens JC, Tanksley SD (1991) Characterization of the level, target sites and inheritance of cytosine methylation in tomato nuclear DNA. Plant Mol Biol 16: 753–770

Moffatt BA, Somerville C (1988) Positive selection for male-sterile mutants of *Arabidopsis* lacking adenine phosphoribosy transferase activity. Plant Physiol 86:1150–1154

Paterson AH, Wing RA (1993) Genome mapping in plants. Curr Opin Biotechnology 4:142–147

Putterill J, Robson F, Lee K, Simon R, Coupland G (1995) The *CONSTANS* gene of *Arabidopsis* promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. Cell 80:847–857

Rick CM (1948) Genetics and development of nine male-sterile tomato mutants. Hilgardia 18:599-633

Rick CM (1980) Tomato Linkage Survey. Report Tomato Genet Coop 30:2–17

Sarfatti M, Katan J, Fluhr R, Zamir D (1989) An RFLP marker in tomato linked to the *Fusarium oxysporum* resistance gene *I2*. Theor Appl Genet 78:755–759

Segal G, Sarfatti M, Schaffer MA, Ori N, Zamir D, Fluhr R (1992) Correlation of genetic and physical structure in the region surrounding the *I2 Fusarium oxysporum* resistance locus in tomato. Mol Gen Genet 231:179–185

Southern E (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98:503–517

Tanksley SD, Ganal MW, Prince JP, De Vicente MC, Bonierbale MW, Broun P, Fulton T, Giovannoni J, Grandillo S, Martin G, Messeguer R, Miller L, Paterson A, Pineda O, Roder M, Wing R, Wu W, Young N (1992) High density molecular linkage maps of the tomato and potato genomes. Genetics 132:1141–1160

Van Daelen RAJJ, Gerbens F, Van Ruissen F, Aarts J, Hontelez J, Zabel P (1993) Long-range physical maps of two loci (*Aps-1* and *GP79*) flanking the root-knot nematode resistance gene (*Mi*) near the centromere of tomato chromosome 6. Pl Mol Biol 23: 185–192

Weide R, Van Wordragen MF, Lankhorst RK, Verkerk R, Hanhart C, Liharska T, Pap E, Stam P, Zabel P, Koornneef M (1993) Integration of the classical and molecular linkage maps of tomato chromosome 6. Genetics 135:1175–1186

Wing RA, Zhang HB, Tanksley SD (1994) Map-based cloning in crop plants. Tomato as a model system: I. Genetic and physical mapping of *jointless*. Mol Gen Genet 242:681–688

Zamir D, Tanksley SD (1988) The tomato genome is comprised largely of fast-evolving, low copy-number sequences. Mol Gen Genet 213:254–261